Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

CRT-2005-12

MODÈLES D'OPTIMISATION POUR LA PLANIFICATION DES SYSTÈMES DE TRANSPORT

par

Teodor Gabriel Crainic
Michael Florian

Mai 2005

Université de Montréal
C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7 Canada
Téléphone : (514) 343-7575 / Télécopieur : (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

HEC ÉCOLE POLYTECHNIQUE MONTRÉAL Université de Montréal

105-05345/

PUBLICATION
CRT-2005-12

MODÈLES D'OPTIMISATION POUR LA PLANIFICATION DES SYSTÈMES DE TRANSPORT

par

Teodor Gabriel Crainic [1,2]
Michael Florian [2]

Ce projet a été financé en partie par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG).

[1] Département de management et technologie, École des sciences de la gestion, Université du Québec à Montréal, C.P. 8888, succursale Centre-ville, Montréal, Canada H3C 3P8

[2] Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7 (mike@crt.umontreal.ca, theo@crt.umontreal.ca)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2005

Abstract

Quantitative methods offer to transportation planning models which predict the demand for transportation of persons or freight within a given region, given the socio-economic characteristics of the population, the industrial profile of the region, and the level of transportation services between the zones of the region. These models aim to predict when and where trips start and for what destinations, the modes used, and the routes followed. The models are then used to compare future scenarios to a reference one in order to select the most appropriate alternative. Operations Research played a major role in the development of models for demand estimation, mode choice, and route selection, as well as of efficient algorithms to address them. This paper presents four important classes of optimization models: spatial interaction models, network equilibrium or trip assignment models, transit route choice models, and multi-mode multi-commodity freight transportation planning models.

Keywords: transportation planning, passenger transportation, freight transportation, models and methods, spatial interaction models, network equilibrium models, transit route choice models, multi-mode multi-commodity freight transportation planning models.

Résumé

Les approches quantitatives à la planification du transport proposent des modèles qui prédisent la demande de déplacements de passagers ou de marchandises dans une région donnée, en se basant sur les caractéristiques socio-économiques de la population, sur le profil industriel de la région et sur les niveaux de service entre l'origine et la destination fournis par l'infrastructure et les services de transport. Le but des modèles est de prédire à quel moment les déplacements débutent, les destinations, les modes utilisés et les routes empruntées. Le processus de planification du transport utilise ces modèles afin de comparer des scénarios futurs à un scénario de référence dans le but d'obtenir des indications quant aux meilleures solutions à adopter. La recherche opérationnelle a joué un grand rôle dans le développement de modèles d'estimation de la demande et de choix de mode et de route, ainsi que d'algorithmes efficaces pour l'obtention de solutions numériques. Quatre catégories importantes de modèles d'optimisation sont présentées dans ce rapport : modèles d'interaction spatiale, modèles d'équilibre de réseau ou d'affectation des déplacements, modèles de choix des routes dans le transport en commun et modèles de planification des réseaux multimodes multiproduits de transport des marchandises.

Mots-clés : planification du transport, transport de passagers, transport de marchandises, modèles et méthodes, modèles d'interaction spatiale, modèles d'équilibre de réseau, choix des routes dans le transport en commun, planification des réseaux multimodes multiproduits de transport des marchandises.

1 Introduction

Les approches quantitatives à la planification du transport proposent des modèles qui prédisent la demande de déplacements de passagers ou de marchandises dans une région donnée, en se basant sur les caractéristiques socio-économiques de la population, sur le profil industriel de la région et sur les niveaux de service entre l'origine et la destination fournis par l'infrastructure et les services de transport. Le but des modèles descriptifs de demande de déplacements est de prédire à quel moment les déplacements débutent, les destinations, les modes utilisés et les routes empruntées.

La théorie et la mise en application des modèles de planification de la demande de transport sont des sujets étoffés, particulièrement concernant les déplacements de passagers dans une zone ou une région urbaine. Les applications aux problèmes de planification des mouvements de marchandises sont plus récentes et s'appuient fortement sur les résultats des travaux réalisés pour le transport des passagers. Dans tous les cas, une grande variété de modèles et de méthodes économétriques et d'optimisation sont utilisés pour formuler et calibrer les modèles à l'aide de données d'enquête. Le processus de planification du transport utilise des modèles descriptifs afin de comparer des scénarios futurs à un scénario de référence dans le but d'obtenir des indications quant aux meilleures solutions à adopter.

Les modèles d'optimisation ont joué un grand rôle dans le développement, depuis les années 70, de modèles d'estimation de la demande, de modèles de choix de mode et de route, ainsi que d'algorithmes efficaces pour l'obtention de solutions numériques. Dans ce chapitre, quatre catégories de modèles d'optimisation seront présentées : modèles d'interaction spatiale (ou modèles de répartition de la demande de déplacements), modèles d'équilibre de réseau (modèles d'affectation des déplacements - « trip assignment »), modèles de choix des routes dans le transport en commun - « transit route choice models » et modèles de planification des réseaux multimodes multiproduits de transport des marchandises.

Ces modèles prennent la forme de problèmes d'optimisation de réseau statique ou d'inégalité variationnelle, même si, en fait, tous les phénomènes de trafic sont temporels. Il faut considérer un temps donné pour lequel la demande de déplacements a été quantifiée. Pour en arriver à une prédiction des flots entre l'origine et la destination et sur les liens (chemins) du réseau, un modèle de réseau est utilisé pour représenter l'offre d'infrastructure et de service de transport. Les modèles d'interaction spatiale sont utilisés pour établir des matrices de demande origine-destination. Les modèles d'équilibre de réseau sont utilisés pour modéliser le choix des routes (« route choice ») sur les réseaux congestionnés, tandis que les modèles de choix des routes dans le transport en commun étudient la fréquence du service offert sur les lignes de transport en commun. Ces modèles font l'objet des quatres prochaines sections.

2 Modèles d'interaction spatiale

Considérons un réseau de transport qui permet l'écoulement d'un seul type de trafic (véhicules ou passagers) sur ses liens. Les nœuds $n \in \mathcal{N}$ représentent les origines, les destinations et les intersections des liens. Les liens $a \in \mathcal{A} \subseteq N \times N$ représentent l'infrastructure ou les services de transport. Si les déplacements qui débutent à l'origine $r \in \mathcal{R} \subseteq \mathcal{N}$ sont représentés par O_r et si les déplacements qui se rendent vers les destinations $s \in \mathcal{S} \subseteq \mathcal{N}$ sont représentés par D_s, la question soulevée est de trouver t_{rs} (ou t_p, où $p = (r, s)$), étant donné le temps et le coût des déplacements u_{rs}.

Le modèle classique, qui est utilisé pour trouver la matrice d'origine-destination t_{rs}, est connu sous le nom de *modèle d'entropie*. La conservation du flot aux origines et aux destinations laisse entendre que

$$\sum_{s \in \mathcal{S}} t_{rs} = O_r, r \in \mathcal{R} \tag{1}$$

et

$$\sum_{r \in \mathcal{R}} t_{rs} = D_s, s \in \mathcal{S}. \tag{2}$$

Bien entendu, la demande pour les déplacements est non négative

$$t_{rs} \geq 0, r \in \mathcal{R}, s \in \mathcal{S}. \tag{3}$$

En l'absence de renseignements complémentaires, nous posons comme principe que la matrice d'origine-destination est la plus vraisemblable, d'où découle la fonction économique

$$\text{Maximiser} \quad -\sum_r \sum_s t_{rs} \ln t_{rs} \tag{4}$$

qui dépend de 1–3. L'objectif est l'interprétation de la maximisation de l'entropie. Le formalisme provient de la théorie de l'information (voir Jaynes [24, 25]). Le modèle a été introduit dans l'analyse du transport et l'analyse régionale par Wilson ([41, 42]) pour les cas où il y a de l'information a priori au sujet de la matrice, par exemple $t^0_{rs}, \forall (r, s)$. Kullback [30] et Snickars et Weibull [37] ont proposé l'utilisation de la fonction

$$\max -\sum_r \sum_s t_{rs} \ln \left(\frac{t_{rs}}{t^0_{rs}} \right). \tag{5}$$

Afin de caractériser la dispersion des déplacements, une contrainte est ajoutée au temps total de déplacement, où C est un temps total de déplacement observé

$$\sum_r \sum_s t_{rs} u_{rs} = C, \tag{6}$$

d'où la fonction économique

$$\min \sum_r \sum_s t_{rs}(\ln t_{rs} + \theta u_{rs}). \tag{7}$$

θ peut être vue comme la variable duale associée à la contrainte 6. Il est évident que, en appliquant les conditions de Karush-Kuhn-Tucker, la solution de 7, qui dépend de 1–3, sera de la forme

$$t_{rs} = \exp(\alpha_r + \beta_s - 1)\exp(-\theta u_{rs}), \forall (r,s) = A_r B_s \exp(-\theta u_{rs}), \tag{8}$$

où α_r et β_s sont les variables duales associées aux contraintes de conservation du flot. Grâce à la convention $t_{rs} \ln t_{rs} = 0$, quand $t_{rs} = 0$, il est possible d'obtenir des solutions à cette classe de problèmes en appliquant un algorithme primal de programmation convexe. Ce problème a la propriété que les variables primales peuvent être exprimées en fonction des variables duales, ce que nous allons maintenant démontrer. Le problème dual lagrangien présenté en 7, qui dépend de 1–2, s'écrit comme suit

$$\begin{aligned} D(\alpha, \beta, t) = \max_{\alpha,\beta} \Bigg\{ & \min_t \sum_r \sum_s t_{rs}(\ln t_{rs} + \theta u_{rs}) \\ & + \sum_r \alpha_r (O_r - \sum_s t_{rs} + \sum_s \beta_s (D_s - \sum_r t_{rs}) \Bigg\}. \end{aligned} \tag{9}$$

En se servant de 8 pour remplacer les variables duales t_{rs}, nous obtenons, en simplifiant

$$D(\alpha, \beta) = \max_{\alpha,\beta} \left\{ \sum_r \alpha_r O_r + \sum_s \beta_s D_s - \sum_r \sum_s \exp(\alpha_r + \beta_s - 1 - \theta u_{rs}) \right\}. \tag{10}$$

Cette propriété a permis l'élaboration d'une procédure de résolution efficace, connue du nom de méthode de balancement, qui est une méthode d'ascension duale pour une variable à la fois. Soit $t^l_{rs}, \alpha^l_r, \beta^l_s$ la solution à une itération l. On obtient alors β^{l+1}_s, qui maximise $D(\alpha^1_r, \beta^{l+1}_s, \beta^l_{s' s' \neq s})$, en annulant la dérivée partielle de la fonction économique par rapport à β_s

$$\frac{\delta D(\beta^{l+1}_s)}{\delta \beta_s} = D_s - \sum_r \exp(-\theta u_{rs} + \beta^{l+1}_s + \alpha^l_r - 1) = 0, \tag{11}$$

qui a comme solution

$$\exp(\beta^{l+1}_s) = D_s / \sum_r \exp(-\theta u_{rs} + \alpha^l_r - 1). \tag{12}$$

En multipliant et en divisant la partie droite de 12 par β^l_s, on obtient

$$\exp(\beta^{l+1}_s) = \frac{D_s}{\sum_r t^l_{rs}} \exp(\beta^l_s). \tag{13}$$

Le ratio $(D_s/\sum_{\alpha} t^l_{rs})$ est désigné sous le nom de facteur de balancement à l'itération l. Puisque la même maximisation unidimensionnelle peut être utilisée pour tous les α_r et β_s, nous obtenons la méthode primale-duale, appelée aussi *méthode de balancement*, décrite ci-après :

Étape 1 (Initialisation) $l = 0, t^0_{rs} = \exp(-\theta u_{rs})\ \forall(r,s)\quad A^0_r = 1, \forall r; B^0_s = 1,\ \forall s$

Étape 2 For $l = 1, 2, ...$

(a) $A^{l+1}_r = \frac{O_r}{\sum_s t^l_{rs}} A^l_r\ \forall r;\quad t^{l+\frac{1}{2}}_{rs} = \frac{O_r}{\sum_s t^l_{rs}} t^l_{rs},\ \forall s$

(b) $B^{l+1}_s = \frac{D_s}{\sum_r t^{l+\frac{1}{2}}_{rs}} B^l_s\ \forall s;\quad t^{l+1}_{rs} = \frac{D_s}{\sum_r t^{l+\frac{1}{2}}_{rs}} t^{l+\frac{1}{2}}_{rs},\ \forall s.$

L'algorithme s'interrompt quand $\|A^{l+1}_r - A^l_r\| \leq \epsilon$ et $\|B^{l+1}_s - B^l_s\| \leq \epsilon$. Rremarquez que $A_r = \ln \alpha_r$ et que la différence entre deux valeurs successives des variables duales est $(\alpha^{l+1}_r - \alpha^l_r) = \ln(O_r/\sum_s t^l_{rs})$.

Cette méthode remonte à au moins 1937, date à laquelle Kruithof [29] l'a utilisée pour prédire la distribution des communications téléphoniques. Demins et Stephan [9] ont redécouvert cette méthode de façon autonome et l'ont appliquée à un problème de classification croisée en statistiques dans le but de simplifier l'ajustement des moindres carrés.

Ce modèle a d'autres propriétés intéressantes. Si $u_{rs} < \infty$ et $\theta > 0$, alors $t_{rs} > 0,\ \forall(r,s)$ et les contraintes de non-négativité 3 ne sont pas requises. En voici l'explication : si t_{rs} est égal à zéro, alors, de 8, soit α_r soit $\beta_s = -\infty$ ce qui suppose que $\sum_s t_{rs} = 0$ ou $\sum_t t_{rs} = 0$, ce qui enfreindrait 1 ou 2.

Une autre variante intéressante de ce modèle se produit lorsque les deux contraintes 1 et 2 sont remplacées par des inégalités. Par exemple, lorsque

$$\sum_s t_{rs} \leq O_r,\ \forall r \tag{14}$$

remplace 1, la solution se trouve en modifiant la méthode de balancement en remplaçant dans le calcul du facteur de balancement dans l'Étape 2(a) $\left(\frac{O_r}{\sum_s t^l_{rs}}\right)$ par $\min\left(1, \frac{O_r}{\sum_s t^l_{rs}}\right)$.

Cette variante du modèle a été étudiée par Jefferson et Scott [26]. Lamond et Stewart [31] ont démontré que la méthode de balancement et ses variantes peuvent être considérées comme un cas spécial de la méthode de projection non orthogonale de Bregman [2], utilisée pour résoudre certains problèmes de programmation convexe. Des variantes de la méthode de balancement ont été examinées par Robillard et Stewart [34], Evans et Kirby [11] et Andersson [1]. L'article de Erlander, Nguyen et Stewart [10] est consacré à la description exhaustive du modèle d'entropie et ses variantes.

3 Modèles et méthodes d'affectation du trafic

Les modèles d'affectation du trafic ont été conçus pour décrire les flots de trafic formés par les usagers d'un réseau de transport, tel qu'un réseau de voies urbaines. Ils peuvent aussi être adaptés pour servir de modèle pour les réseaux aériens, ferroviaires, et autres. On considère que certains attributs du réseau sont connus et que la demande de déplacements est définie par une matrice de demande origine-destination, tel que décrit dans la section précédente, ou est définie par les fonctions de demande.

Dans cette section, les plus importants modèles d'affectation du trafic sont décrits et quelques algorithmes élémentaires sont proposés pour les résoudre. Les preuves des résultats proposés, l'analyse de la convergence des algorithmes et une exposition détaillée des autres modèles sont proposés dans, par exemple, Sheffi [36], Patricksson [33] et Florian et Hearn [13].

Pour simplifier la notation, nous considérons un modèle de réseau de transport qui comporte un seul type de flot de véhicules sur les liens dirigés du réseau. Les nœuds $i, i \in \mathcal{N}$, représentent les origines, les destinations et les intersections ; les arcs $a, a \in \mathcal{A}$, représentent les liens de transport. Les demandes origine-destination (O-D) engendrent des flots sur les liens $v_a, a \in \mathcal{A}$; le coût du déplacement sur un lien est fourni par une fonction de coût usager $s_a(v)$ où v est le vecteur $(v_a)_{a\in\mathcal{A}}$ des flots sur tous les liens du réseau. Les fonctions de coût modélisent les temps et délais sur les liens ou des coûts divers tels que les péages et la consommation de carburant, considérés non négatifs. Soit $\mathcal{P}$ l'ensemble des paires O-D, $\mathcal{K}_p$, $p \in \mathcal{P}$, l'ensemble des chemins orientés qui relient la paire p et $\mathcal{K}$, l'ensemble de tous les chemins. La demande de transport t_p pour la paire p utilise des chemins orientés et les flots h_k sur les chemins obéissent aux contraintes de conservation du flot et de non-négativité :

$$\sum_{k\in\mathcal{K}_p} h_k = t_p \quad \forall p \in \mathcal{P}, \tag{15}$$

$$h_k \geq 0 \qquad \forall k \in \mathcal{K}. \tag{16}$$

Les flots sur les liens sont donnés par

$$v_a = \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{K}_p} \delta_{ak} h_k \quad \forall a \in \mathcal{A} \tag{17}$$

où $\delta_{ak} = 1$ si le lien appartient au chemin k et est égal à zéro autrement.

Définissons $\Delta = (\delta_{ak})$ comme étant la matrice d'incidence arc-chemin $|\mathcal{A}| \times |\mathcal{K}|$ telle que $v = \Delta h$, où h est le vecteur $(h_k)_{k\in\mathcal{K}}$ des flots sur les chemins pour toutes les paires O-D. Le coût $s_k(= s_k(h))$ de chaque chemin k s'écrit

$$s_k = \sum_{a\in\mathcal{A}} \delta_{ak} s_a(v) = \sum_{a\in\mathcal{A}} \delta_{ak} s_a(\Delta h), \quad \forall k \in \mathcal{K}_p, p \in \mathcal{P}, \tag{18}$$

et $u_p(= u_p(h))$ est, par définition, le coût du chemin de moindre coût pour la paire O-D p

$$u_p = \min_{k \in \mathcal{K}_p} \; s_k, \quad \forall p \in \mathcal{P}. \tag{19}$$

Pour chaque $p \in \mathcal{P}$, soit nous obtenons la demande de déplacements t_p d'une matrice de demande fixe O-D, dans quel cas $t_p = \bar{t}_p$, ou encore elle est fournie par la fonction de demande $t_p(u)$, où u est le vecteur des valeurs des déplacements de moindre coût, $(u_p)_{p \in \mathcal{P}}$, pour toutes les paires O-D du réseau :

$$t_p = t_p(u) \qquad \forall p \in \mathcal{P}. \tag{20}$$

3.1 Modèles d'optimisation système et d'optimisation usager

Les modèles d'affectation du trafic basés sur l'« optimisation système » supposent que les déplacements sur le réseau suivent des chemins de telle manière que l'utilisation du réseau se fait dans « l'intérêt commun ». Si les demandes t_p sont fixes, l'objectif est de satisfaire un principe normatif selon lequel le coût (ou le temps) moyen des déplacements doit être minimisé. Puisque la demande totale est une constante, il s'agit de l'équivalent de la minimisation du coût total sur le système. Le modèle d'optimisation système à demande fixe se présente comme suit :

$$\text{Minimiser} \quad \sum_{a \in \mathcal{A}} s_a(v) v_a \tag{21}$$

sous les contraintes 16 et 17 et $t_p = \bar{t}_p$.

Si, cependant, la demande de déplacements est élastique, c'est-à-dire dictée par les fonctions de demande 20, le modèle d'optimisation système tente de maximiser les avantages économiques nets pour les usagers du réseau. Les principes économiques établis nous permettent de savoir que l'avantage pour les usagers, pour n'importe quelle paire O-D $p \in \mathcal{P}$, s'obtient en calculant l'aire sous la courbe de demande $t_p(u)$. On suppose que cette fonction a une fonction inverse $w_p(t_p) = u_p$. L'avantage économique peut donc s'écrire $\int_0^{t_p} w_p(y)dy$. Par conséquent, dans ce cas, le modèle d'optimisation système se décrit comme suit :

$$\text{Maximiser} \quad \sum_p \int_0^{t_p} w_p(y)dy - \sum_{a \in \mathcal{A}} s_a(v) v_a \tag{22}$$

sous les contraintes 15, 16, 17 et 20.

Les problèmes rencontrés dans les modèles d'optimisation système apparaissent aussi dans les modèles d'optimisation usager qui, eux, tentent de décrire avec plus de précision la répartition des usagers sur le réseau sachant qu'il est impossible pour un usager

d'améliorer à lui seul ses coûts de déplacement. Par conséquent, les modèles descriptifs des flots de trafic supposent que les usagers se trouvent en équilibre de Wardrop [40], un cas spécial d'équilibre de Nash. Le problème se décrit comme suit : Soit h^* et u^* de façon à satisfaire les conditions suivantes :

$$(s_k(h^*) - u_p^*)h_k^* = 0 \qquad \forall k \in K_p, p \in \mathcal{P} \tag{23}$$

$$s_k(h^*) - u_p^* \geq 0 \qquad \forall k \in K_p, p \in \mathcal{P} \tag{24}$$

$$\sum_{k \in K_p} h_k^* - t_p = 0 \qquad \forall p \in \mathcal{P} \tag{25}$$

$$h^* \geq 0, u^* \geq 0, \tag{26}$$

où $t_p = \bar{t}_p$ est utilisé lorsque la demande est fixe et $t = t(u^*)$ est utilisé lorsque la demande est élastique. Les flots d'équilibre sur les liens, v^*, peuvent être calculés à partir des flots sur les chemins h^* en appliquant 17.

Les deux premières conditions visent à ce que, pour tout $p \in \mathcal{P}$, seuls les chemins de moindre coût soient utilisés. La troisième condition établit un rapport entre le total des flots sur les chemins et la demande totale, compte tenu des coûts minimum sur les chemins. Cet énoncé général du problème est connu sous le nom de *modèle d'équilibre de réseau*. Ce modèle est utilisé dans de nombreuses applications, dont les réseaux électriques, les réseaux de conduites d'eau et les problèmes d'équilibre de prix spatial. Florian et Hearn [13] fournissent des exemples des applications citées. Ils résument également les résultats des projets de recherche de nombreux auteurs, démontrant ainsi que l'analyse mathématique du modèle d'équilibre de réseau peut être facilitée par sa reformulation soit en problème de complémentarité non linéaire, soit en problème de point fixe ou encore en problème d'inégalité variationnelle.

Par contre, les reformulations du modèle d'équilibre de réseau utilisées dans la planification du transport sont des problèmes d'optimisation. Les hypothèses principales prévoient que les fonctions de coût et de demande peuvent être séparées, c'est-à-dire qu'elles ont la forme $s_a(v) = s_a(v_a)$ et $t_p(u) = t_p(u_p)$. En d'autres termes, le coût sur un lien dépend uniquement du flot sur le lien, et la demande pour la paire O-D p est fonction uniquement du temps de déplacement minimum pour cette paire O-D. On considère, de plus, que les fonctions de coût sont convexes et que les fonctions de demande sont strictement monotones. En tenant compte de ces conditions, le problème d'optimisation usager de la demande élastique peut être formulé comme le programme convexe

$$\min \sum_{a \in A} \int_0^{v_a} s_a(x) d_x - \sum_{p \in P} \int_0^{t_p} w_p(y) dy \tag{27}$$

compte tenu de 15, 16, 17 et 20. Quand la demande est fixe, le problème du côté usager devient

$$\min \sum_{a \in \mathcal{A}} \int_0^{v_a} s_a(x) dx \tag{28}$$

compte tenu de 15, 16 et 17 et $t_p = \bar{t}_p$.

On constate que les solutions à ces problèmes sont équivalentes à ce qui est proposé en 23–26 pour la demande fixe et la demande élastique. Cette situation découle directement des conditions Karush-Kuhn-Tucker des deux problèmes. Le lien entre les conditions du modèle d'équilibre réseau et les modèles d'optimisation système se reflète également dans leurs conditions Karush-Kuhn-Tucker. Rien de plus simple que de vérifier qu'ils ont la même forme, les termes $s_k(h^*)$ étant calculés à partir des coûts marginaux sur les liens, $s_a(v_a^*) + s_a^{'}(v_a^*)v_a^*$. Le lien significatif qui les relie est donc que la solution au modèle d'optimisation système se trouve en équilibre en ce qui concerne les coûts marginaux, tandis que le modèle d'optimisation usager est en équilibre en ce qui concerne les coûts moyens.

3.2 Algorithmes pour l'affectation du trafic pour le modèle d'optimisation usager

Trois méthodes de base seront présentées dans cette section : deux pour le modèle d'optimisation usager à demande fixe et une pour le modèle d'optimisation usager à demande élastique. Puisque les deux modèles d'optimisation (usager et système) ont comme seule différence les termes de la fonction économique qui contiennent les fonctions de coût sur les liens, les algorithmes conçus pour ces deux modèles pourront être utilisés pour les modèles équivalents d'optimisation système. La modification principale qui doit être apportée se trouve dans les dérivées des fonctions économiques, qui elles ont une incidence sur les coûts des plus courts chemins. Dans les modèles « usager » on les obtient de $s_a(v_a)$, mais dans les modèles « système » , on les obtient de $s_a(v_a) + s_a^{'}(v_a)v_a$. Cet ajustement correspond à l'observation faite précédemment concernant les conditions Karush-Kuhn-Tucker des modèles.

3.2.1 Décomposition cyclique des O-D avec équilibre des chemins

Les modèles « usager » rendent compte des flots et des coûts sur les chemins pour chaque paire O-D. La solution la plus intuitive se trouve donc au niveau des flots sur les chemins avec décomposition extérieure par paire O-D. Cette approche, identique à la décomposition (ou relaxation) de Gauss-Seidel, est également connue sous le nom de décomposition cyclique, puisqu'un seul problème « affectation usager » est résolu pour chaque paire O-D durant un cycle de l'algorithme, les flots pour les autres paires O-D demeurant fixes. La stratégie cyclique générale est présentée ci-dessous :

Étape 0. Soit un ensemble de solutions initiales $p = 0, p' = 0$.

Étape 1. Si $p' = |P|$, le nombre total de paires O-D, STOP ;
Autrement poser $p = p \bmod |P| + 1$ et continuer ;

Étape 2. Si la solution actuelle est optimale pour le sous-problème p (29–33), alors $p' = p' + 1$ et revenir à l'étape 1 ;
Autrement, résoudre le sous-problème p, mettre à jour les flots, faire $p' = 0$ et revenir à l'étape 1.

Voici un énoncé du sous-problème lorsqu'il est appliqué au problème « usager » avec demande fixe :

$$\min \sum_{a \in \mathcal{A}} \int_0^{v_a^p + \overline{v}_a} s_a(x) dx \tag{29}$$

compte tenu de

$$\sum_{k \in \mathcal{K}_p} h_k = \overline{t}_p, \tag{30}$$

$$h_k \geq 0, \quad \forall k \in \mathcal{K}_p, \tag{31}$$

où

$$v_a = \sum_{p \neq p} \sum_{k \in \mathcal{K}_p} \delta_{ak} h_k \tag{32}$$

et

$$v_a^p = \sum_{k \in \mathcal{K}_p} \delta_{ak} h_k, \quad \forall a \in \mathcal{A}. \tag{33}$$

Les algorithmes d'équilibre de chemin utilisés pour résoudre 29–33 se déroulent dans l'espace des flots sur les chemins et proposent des solutions où tous les chemins ont le même coût. Le plus simple de ces algorithmes trouve le plus court chemin et le plus long chemin et fait passer les flots de l'un à l'autre pour égaliser les coûts. Soit $\mathcal{K}_p^+ = \{k \in \mathcal{K}_p \mid h_k > 0\}$, l'ensemble des chemins ayant des flots positifs, l'*algorithme d'équilibre de chemin* se présente comme suit :

Étape 0. Soit une solution initiale v_a^p; $s_a = s_a(v_a^p + \overline{v}_a)$ et $\mathcal{K}_p^+$ initial.

Étape 1. Calculer les coûts des chemins présentement utilisés : $s_k, k \in \mathcal{K}_p^+$. Trouver k_1 de sorte que $s_{k_1} = \min_{k \in \mathcal{K}_p^+} \{s_k\}$ et k_2 de sorte que $s_{k_2} = \max_{k \in \mathcal{K}_p^+} \{s_k\}$. Si $(s_{k_2} - s_{k_1}) \leq \epsilon$, aller à l'étape 4, autrement procéder à l'étape 2.

Étape 2. Redistribuer le flot du chemin k_2 sur le chemin k_1 jusqu'à ce qu'il soit réduit à zéro ou que les coûts sur les deux chemins soient égaux.

Étape 3. Compte tenu des nouveaux flots sur les chemins h_{k_1} et h_{k_2}, mettre à jour les flots sur les liens v_a^p, les coûts sur les liens s_a et les coûts sur les chemins s_k.

Étape 4. Calculer le plus court chemin $\bar{k}$ au coût $s_k = \min_{k \in \mathcal{K}_p} \{s_k\}$; si $\bar{s}_k < \min_{k \in K_p^+} \{s_k\}$, alors le chemin $\bar{k}$ est ajouté à l'ensemble des chemins conservés, $\mathcal{K}_p^+ = \mathcal{K}_p^+ \cup \bar{k}$, retourner à l'étape 1 ; autrement STOP.

L'algorithme présenté ci-dessus n'est que l'un des systèmes possibles d'équilibre des chemins. Par exemple, l'adaptation, pour le sous-problème, des algorithmes du gradient réduit ou du gradient projeté entraîne une étape d'équilibre sur tous les chemins

$k, k \in \mathcal{K}_p^+$. Par ailleurs, il n'est pas nécessaire de résoudre à optimalité chacun des sous-problèmes ; une ou deux étapes de l'équilibration peuvent être exécutées pour chacune des paires O-D.

Cet algorithme convient au problème « usager » avec demande fixe lorsque le nombre de chemins utilisés pour une paire O-D est plus ou moins petit, comme dans un réseau aérien ou ferroviaire, et lorsque le nombre de paires O-D est plus ou moins petit également. Dans de nombreux modèles de réseaux routiers, les chemins et les flots exigent des ressources de stockage de plus en plus considérables. Pour cette raison, les méthodes décrites ci-dessous sont davantage utilisées puisqu'elles se déroulent dans l'espace des flots sur les liens.

3.2.2 Méthode d'approximation linéaire

La méthode d'approximation linéaire propose l'un des algorithmes les plus simples pour minimiser une fonction convexe sous contraintes linéaires. L'adaptation de cet algorithme pour résoudre le modèle d'optimisation système et le modèle d'optimisation usager produit des méthodes qui n'exigent que le calcul des plus courts chemins et la minimisation unidimensionnelle d'une fonction convexe.

Nous allons d'abord examiner l'adaptation au problème d'optimisation usager avec demande fixe. Prenant comme point de départ une solution réalisable, la méthode produit une direction de descente réalisable en fournissant la solution à un sous-problème que l'on obtient par la linéarisation de la fonction économique. Ensuite, une solution améliorée provient du segment de ligne, entre la solution actuelle et la solution du sous-problème. Soit $S(v)$ la fonction économique. L'approximation linéarisée de cette fonction à une itération intermédiaire l, lorsque la solution courante est v^l, s'écrit :

$$S(v^l) + \nabla S(v^l)(y - v^l). \tag{34}$$

Puisque $S(v^l)$ et $\nabla S(v^l)v^l$ sont des constantes, le sous-problème linéarisé qu'il faut résoudre se réduit à

$$\min \sum_{p \in \mathcal{P}} \sum_{k \in \mathcal{K}_p} \sum_{a \in A} s_a(v_a^l)\delta_{ak} h_k \tag{35}$$

sous les contraintes

$$\sum_{k \in \mathcal{K}_p} h_k = \bar{t}_p \qquad \forall p \in \mathcal{P}, \tag{36}$$

$$h_k \geq 0 \qquad \forall k \in \mathcal{K}. \tag{37}$$

En changeant l'ordre de la sommation en 35 et en utilisant 18, nous obtenons

$$\min \sum_{p \in \mathcal{P}} \sum_{k \in \mathcal{K}_p} s_k^l h_k. \tag{38}$$

Puisque les termes de la fonction économique 38 peuvent être séparés par paire O-D p, on obtient la solution du sous-problème linéarisé en calculant les plus courts chemins pour chacune des paires O-D p et en affectant la demande $\bar{t}_p$ aux liens sur les chemins. Une telle allocation ou affectation est connue sous le nom d'affectation *tout ou rien*, qui produit le vecteur de flot sur l'arc

$$y_a^l = \sum_{k \in \mathcal{K}} \delta_{ak} h_k^l \qquad \forall a \in \mathcal{A} \tag{39}$$

et la direction de la descente

$$d_a^l = (y_a^l - v_a^l) \qquad \forall a \in \mathcal{A}. \tag{40}$$

Une itération de l'algorithme d'approximation linéaire est réalisée en effectuant une recherche linéaire entre v^l et y^l. Il s'agit là d'un problème avec une seule variable

$$\min_{0 \leq \lambda \leq 1} S(v^l + \lambda d^l). \tag{41}$$

L'adaptation de la *méthode de l'approximation linéaire* produit l'algorithme suivant :

Étape 0. Trouver une solution initiale v^l; $s^l = s(v^l)$; $l = 1$.

Étape 1. Réaliser une affectation tout ou rien en tenant compte des coûts actuels sur l'arc $s(v^l)$ pour obtenir le vecteur de flot sur l'arc y^l ; Soit $d^l = (y^l - v^l)$.

Étape 2. Vérifier si un critère d'arrêt établi au préalable a été satisfait. Si oui, arrêter ; autrement procéder à

Étape 3. Trouver la taille optimale du pas λ^l en trouvant la solution à 41.

Étape 4. Mettre à jour les flots sur l'arc $v^{l+1} = v^l + \lambda^l d^l$ et les coûts sur l'arc $s^{l+1} = s(v^{l+1})$; établir $l = l + 1$ et revenir à l'étape 1.

L'algorithme présente de nombreux avantages pour la solution du modèle d'équilibre du réseau symétrique à demande fixe. Les chemins utilisés pour calculer la direction de la descente à une itération donnée sont produits au besoin et ne sont pas obligatoirement conservés dans des itérations successives. Les ressources requises pour le stockage sont donc modestes et n'augmentent pas selon le nombre d'itérations. Pour chaque itération, seuls les flots v^l et les coûts sur les liens s^l sont conservés, en plus des données sur les liens du réseau.

Puisque $S(v)$ est une fonction convexe et $\nabla S(v) = s(v)$, pour toute itération nous avons

$$S(v^*) \geq S(v^l) + s(v^l)(y^l - v^l). \tag{42}$$

Ainsi, la meilleure borne inférieure atteignable jusqu'à l'itération l est

$$BLB = \max_{i=1:..l} S(v^l) + s\left(v^l\right)\left(y^l - v^l\right) \tag{43}$$

et donc, un critère naturel d'arrêt survient lorsque l'erreur relative

$$\frac{S(v^l) - BLB}{S(v^l)} \tag{44}$$

est plus petite qu'un paramètre prédéterminé.

Notons que cet algorithme propose l'interprétation intuitive selon laquelle les usagers modifient leur choix de route en fonction de la congestion. Cet algorithme se rapproche donc des algorithmes heuristiques qui ont été proposés et utilisés pour résoudre ce problème. Par contre, le désavantage principal de la méthode d'approximation linéaire, c'est qu'elle produit parfois une lente convergence près de la solution optimale parce que son taux asymptotique de convergence est arithmétique. En raison de cette situation, de nombreuses variantes de cet algorithme ont été développées qui tentent d'améliorer son taux de convergence. Deux des algorithmes qui ont le mieux réussi sont fondés sur la méthode PARTAN et sur la décomposition simpliciale : ils sont tous deux examinés dans l'étude d'ensemble effectuée par Florian et Hearn [13].

Le problème d'optimisation usager avec demande variable peut être résolu par une méthode d'approximation linéaire partielle qui effectue une linéarisation sur quelques variables seulement de la fonction économique. Étant donné la solution (v^l, t^l) à l'itération l, le but du sous-problème est de trouver (y^l, z^l) afin qu'une recherche linéaire puisse être effectuée à la fois sur le flot sur le lien et la demande O-D. Dans le cas présent, il est naturel de linéariser uniquement les fonctions de coût de l'arc. Soit $W(z) = \sum_p \int_0^{z_p} w_p(x)dx$, le sous-problème qui en découle à l'itération l est

$$\text{Minimiser} \quad \sum_{p\in\mathcal{P}} \sum_{k\in\mathcal{K}} \sum_{a\in A} s_a(v_a^l)\delta_{ak}h_k - W(z) \tag{45}$$

sous contraintes

$$\sum_{k\in\mathcal{K}_p} h_k - z_p = 0 \qquad \forall p \in \mathcal{P}, \tag{46}$$

$$h_k \geq 0 \quad \forall k \in \mathcal{K}, z_p \geq 0 \qquad \forall p \in \mathcal{P}. \tag{47}$$

Pour trouver la solution à ce sous-problème, il faut d'abord établir que $u_p^l, p \in \mathcal{P}$, représente les coûts des chemins les plus courts basés sur les coûts actuels sur les liens $s(v^l)$; il faut ensuite simplifier 45 en utilisant 46 et 47 afin d'obtenir

$$\text{Minimiser} \quad \sum_{p\in\mathcal{P}} (u_p^l z_p - W_p(z_p)) \tag{48}$$

compte tenu de

$$z_p \geq 0 \quad \forall p \in \mathcal{P}. \tag{49}$$

À partir des conditions d'optimalité du sous-problème, les z_p^l sont établis analytiquement comme suit :

$$z_p^l = \begin{cases} z_p(u_p^l) & if \;\; t_p(u_p^l) \geq 0, \\ 0 & \text{autrement.} \end{cases} \tag{50}$$

La recherche linéaire entre (v^l, t^l) et (y^l, z^l) s'ensuit.

Les critères de convergence pour la méthode d'approximation linéaire partielle peuvent être basés sur la meilleure borne inférieure ou sur le nombre maximal d'itérations. En particulier, la borne inférieure fournie par l'approximation linéaire partielle pour le problème d'optimisation usager avec demande variable est

$$BLB = \max_l S(v^l) + s(v^l)(y^l - v^l) - W(t^l) - w(t^l)(z^l - t^l). \tag{51}$$

3.3 Le problème de l'usager comme inégalité variationnelle

Le succès des algorithmes d'optimisation dans la résolution des applications du modèle d'équilibre de réseau avec fonctions de demande et de coût séparables a motivé les chercheurs à trouver des formulations pour le modèle général d'équilibre de réseau, modèle où la séparation n'est pas une condition posée d'avance. Tel que mentionné précédemment, des formulations de point fixe et de complémentarité non-linéaire ont aussi été élaborées. Cependant, l'intérêt pour les modèles d'application générale porte surtout sur la formulation du problème en tant qu'inégalité variationnelle (IV). Dans cette section, nous allons présenter certains résultats obtenus pour le problème de l'usager à demande fixe. Des extensions au problème de demande élastique se trouvent dans les références, voir entre autres Florian et Hearn [13].

Le résultat théorique fondamental stipule que v^* fournit la solution au problème de demande fixe pour le modèle d'équilibre de réseau si et seulement si il permet de trouver la solution au problème d'IV suivant :

$$\sum_{a \in \mathcal{A}} s_a(v^*)(v_a - v_a^*) \geq 0 \tag{52}$$

pour tout (v, t), tel que

$$\sum_{k \in \mathcal{K}_p} h_k = \bar{t}_p, \qquad \forall p \in \mathcal{P}, \tag{53}$$

$$h_k \geq 0, \qquad \forall k \in \mathcal{K}, \tag{54}$$

$$v_a = \sum_{k \in \mathcal{K}} \delta_{ak} h_k, \qquad \forall a \in A. \tag{55}$$

Nous appelons ce modèle, le modèle IV de l'usager. Pour simplifier la description des deux algorithmes ci-dessous, nous en comprimons la notation en définissant Θ comme étant l'ensemble des flots réalisables en fonction du vecteur de flot sur l'arc v : Trouver v^* dans Θ de sorte que

$$s(v^*)(v - v^*) \geq 0, \qquad \forall v \in \Theta, \tag{56}$$

où

$$\Theta = \left\{ v : v_a = \sum_{k \in \mathcal{K}} \delta_{ak} h_k, \sum_{k \in \mathcal{K}_p} h_k = \bar{t}_p, \forall p \in \mathcal{P}, a \in A \right\}.$$

3.3.1 Méthode de projection pour IVU-DF

Une méthode répandue pour résoudre ce modèle est fondée sur la projection successive dans l'espace des flots sur les liens Θ. Il est à noter qu'une autre solution consiste en la décomposition cyclique O-D (méthode décrite précédemment) avec utilisation d'une méthode de projection pour résoudre le sous-problème qui en résulte pour chaque paire O-D dans l'espace des flots sur les chemins.

Soit Q, une matrice symétrique positive définie, et soit $\bar{v}, \bar{v} \in \Theta$, un flot réalisable. Une nouvelle fonction de coût $\hat{s}(v)$ est définie sur Θ comme suit :

$$\hat{s}(v) = Qv + c, \tag{57}$$

où

$$c = \rho s(\bar{v}) - Q\bar{v} \tag{58}$$

et ρ est une constante positive. Le choix des valeurs appropriées de ρ sera étudié ci-dessous. Puisque $\hat{s}(v)$ est linéaire et possède une jacobienne symétrique, la solution $\hat{v}$ de

$$\hat{s}(\hat{v})(v - v) \geq 0 \qquad \forall v \in \Theta \tag{59}$$

est fournie par la solution d'un modèle symétrique d'équilibre de réseau, obtenue par l'utilisation d'un des algorithmes de la section précédente. L'inégalité 59 est identique à

$$(Q\bar{v} = \rho s(\bar{v}) - Q\bar{v})(v - \bar{v}) \geq 0 \qquad \forall v \in \Theta \tag{60}$$

ce qui définit l'application M_ρ sur Θ, où $\bar{v}$ est intégré en $\hat{v}$, la solution unique de 60. Ceci signifie que chaque point fixe sur M_ρ est la solution de 56. Pour le confirmer, soit $\bar{v}$, un point fixe, c'est-à-dire $M_\rho(\bar{v}) = \bar{v}$. En établissant $\hat{v} = \bar{v}$ en 60, il en résulte $\rho s(\bar{v})(v - \bar{v}) \geq 0 \; \forall v \in \Theta$; en divisant par $\rho > 0$, on constate que, effectivement, $\bar{v}$ est la solution de 56.

L'algorithme qui en résulte constitue une méthode de projection puisque $\hat{v}$ est la projection du point $\bar{v} - \rho Q^{-1} s(\bar{v})$ sur l'ensemble Θ, où la projection est définie par rapport à la norme Q, c'est-à-dire $v = \mathrm{Proj}_{Q,\Theta}(\bar{v} - \rho Q^{-1} s(\bar{v}))$. On obtient donc l'*algorithme de projection* suivant :

Étape 0. Trouver une solution initiale v^l. Soit $s(v^l) = Qv^l + c$, $l = 1$.

Étape 1. Calculer v^{l+1} en tant que $\mathrm{PROJ}_{Q,\Theta}(v^l - pQ^{-1}s(v^l))$ ou, de façon équivalente, résoudre

$$\text{Minimiser} \quad (v - v^l)\hat{s}(v^l) + \frac{1}{2\rho}(v - v^l)Q(v - v^l) \quad \text{sujet à} \quad v \in \Theta.$$

Étape 2. Si $\|v^{l+1} - v^l\| \leq \epsilon$, STOP. Autrement, soit $l = l + 1; \hat{s}(s^{l+1}) = Qv^{l+1} + c$ et revenir à l'étape 1.

La convergence de l'algorithme prévoit que s doit être continuellement dérivable et fortement monotone, c'est-à-dire

$$(s(v_1) - s(v_2))(v_1 - v_2) \geq \alpha |v_1 - v_2|^2 \qquad \forall v_1, v_2 \in \Theta, \tag{61}$$

où $\alpha > 0$ est le module d'une forte monotonie, pourvu que ρ soit suffisamment petit. Cette condition technique pourrait s'avérer difficile à vérifier dans la pratique puisque ρ dépend à la fois de α et des valeurs propres (plus de Θ) de la matrice $\nabla s(v) Q^{-1} \nabla s(v)$.

3.3.2 Décomposition simpliciale pour IVU-DF

Le succès de la méthode d'approximation linéaire pour les modèles séparables à demande fixe de grande taille provient du fait qu'elle se déroule dans l'espace des flots sur les liens et qu'elle passe alternativement des calculs des plus courts chemins aux recherches linéaires. Il faut de nombreuses modifications pour généraliser la méthode pour utilisation avec IVU-DF. D'abord, en l'absence d'une fonction économique, il faut utiliser une fonction auxiliaire pour produire les coûts pour les calculs des chemins les plus courts et pour le contrôle du progrès de la méthode. Ensuite, les simples recherches linéaires doivent être remplacées par la résolution des sous-problèmes d'IV sur des simplices.

La fonction auxiliaire G est appelée fonction d'écart. Elle est définie comme suit : pour tout $\tilde{v} \in \Theta$,

$$G(\tilde{v}) = \max_{v \in \Theta} \quad s(\tilde{v})(\tilde{v} - v).$$

Notons que $G(\tilde{v}) \geq 0$ pour tout $\tilde{v} \in \Theta$. Par ailleurs, si $\tilde{v}$ ne calcule pas la valeur de 56, $G(\tilde{v}) > 0$, et $G(\tilde{v}) = 0$ si et seulement si $\tilde{v} = v^*$, la solution optimale. Trouver la valeur de 56 est équivalent à trouver un v^* qui calcule la valeur de

$$\min_{\tilde{v} \in \Theta} G(\tilde{v}). \tag{62}$$

On peut réécrire cette équation comme suit

$$\min_{\tilde{v} \in \Theta} \left\{ \max_{v \in \Theta} \{ s(\tilde{v})(\tilde{v} - v) \} \right\}, \tag{63}$$

où, pour chaque $\tilde{v}$, le problème intérieur de maximisation est équivalent au sous-problème de programmation linéaire (plus court chemin) de la méthode d'approximation linéaire.

Dans le contexte de l'affectation d'équilibre du trafic, $G(\tilde{v})$ propose une interprétation importante qui nous permet de déduire 62 intuitivement. Pour ce faire, notons que

$$G(v) = s(\tilde{v})\tilde{v} + \max_{v \in \Theta} -s(\tilde{v})v \tag{64}$$

$$= s(v)v - \min_{v \in \Theta} s(\tilde{v})v \tag{65}$$

$$= \sum_{a \in \mathcal{A}} s_a(\tilde{v}_a)\tilde{v}_a - \sum_{p \in \mathcal{P}} \tilde{u}_p \bar{t}_p, \tag{66}$$

où $\tilde{u}_p = u_p(\tilde{v})$ correspond, par définition 19, au chemin le plus court, étant donné $\tilde{v}$, pour la paire O-D p. Le premier terme de l'expression en 66 représente le coût total du système et le deuxième terme représente le coût total du système si toute la demande est acheminée sur les chemins les plus courts en tenant compte de $\tilde{v}$. Manifestement, ces deux termes sont égaux si et seulement si $\tilde{v} = v^*$, vecteur de flot optimal des usagers.

L'algorithme de *décomposition simpliciale* s'écrit comme suit :

Étape 0. Soit Θ' un ensemble qui initialement contient un seul point extrême de Θ. Considérons $\delta > 0$ donné. Fixer $\bar{G}^l = \infty$ et $l = 1$. Soit $\{\epsilon_l\}$, une séquence positive monotone strictement décroissante qui converge vers zéro.

Étape 1. Trouver v^l dans l'enveloppe convexe de (Θ^l) tel que $s(v^l)(v - v^l) \geq \epsilon_l$, pour chaque v dans l'enveloppe convexe de (Θ^l). Soit $\mathcal{D}^l$, l'ensemble des éléments de Θ^l de poids zéro dans l'expression de v^l comme combinaison convexe des points extrêmes de Θ^l.

Étape 2. Résoudre

$$\text{Minimiser} \quad s(v^l)y \quad \text{sujet à} \quad y \in \Theta.$$

Dénoter la solution du programme linéaire psr y^l. Si $G(v^l) = s(v^l)(v^l - y^l) = 0$, STOP. Autrement

(i) si $G(v^l) \geq \bar{G}^l - \delta$, soit $\Theta^{l+1} = \Theta^l \cup \{y^l\}$,

(ii) si $G(v^l) < \bar{G}^l - \delta$, soit $\Theta^{l+1} = (\Theta^l - D^l) \cup \{y^l\}$.

Soit $\bar{G}^{l+1} = \min\{\bar{G}^l, G(v^l)\}$. Fixer $l = 1 + 1$ etller à l'étape 1.

Il est à noter que l'étape 1 comporte une inégalité variationnelle « relaxée », c'est-à-dire qu'une solution approximative est suffisante. Toutes les méthodes peuvent être utilisées, mais une méthode de projection est habituellement utilisée parce que le domaine est un simplexe. L'étape 2 présente le calcul habituel des chemins les plus courts, comme dans la méthode de l'approximation linéaire. Puisque $\{G(v^l)\}$ n'est pas monotone, $\bar{G}^l$ a été introduit pour obtenir une séquence monotone décroissante, ce qui permet de prouver la convergence de la méthode.

4 Modèles de choix de routes dans le transport en commun

Les modèles de choix de routes dans le transport en commun ou modèles d'affectation dans le transport en commun ont comme objectif de décrire les flots sur un réseau de lignes de transport en commun où l'horaire est fixe. Ce qui les distingue des modèles

et des méthodes d'affectation de trafic présentés dans la section 3, c'est l'attente. : les usagers du transport en commun doivent attendre le premier véhicule (autobus) de la ligne choisie. De plus, l'accès au point d'arrêt sous-entend un temps d'attente : le temps de marche à l'arrêt, les correspondances (s'il y a lieu), et le temps à bord du véhicule. L'exposition qui suit s'inspire du travail effectué par Spiess et Florian [39].

Considérons un réseau de transport en commun qui contient un ensemble de nœuds, de lignes de transport en commun (chacune étant définie par une liste de nœuds ordonnés où les montées à bord et les descentes sont permises), et de liens de marche (chacun défini par deux nœuds). Les temps associés aux liens de marche et aux segments des lignes de transport en commun sont constants. La distribution de l'intervalle entre les arrivées des véhicules est connue pour chacune des lignes qui desservent les nœuds qui se trouvent sur l'itinéraire d'une ligne de transport en commun. Par conséquent, il est possible de calculer le temps combiné de l'arrivée prévue du premier véhicule, et ce, pour tous les sous-ensembles de lignes rattachées à un nœud, ainsi que la probabilité que l'une des lignes arrivera la première.

Avant de formuler le modèle mathématique qui correspond à la sélection du choix de ligne de transport en commun, il faut préciser qu'un lien de marche peut être remplacé (théoriquement) par une ligne de transport en commun d'un lien ayant un temps d'attente de zéro (fréquence infinie). Par ailleurs, nous supposons que le réseau sous-jacent est fortement connecté. L'objectif est de minimiser les temps d'attente et de déplacement (statistiquement) *espérés* ou, encore, le coût total généralisé *espéré* si les temps d'attente et les temps de déplacement ont des pondérations différentes (c'est-à-dire qu'attendre est plus pénible qu'être dans le véhicule).

Le réseau est formé de quatre types d'arcs : les arcs d'attente (aucun temps de déplacement), les arcs à bord (pas d'attente), les arcs de descente (pas de temps de déplacement, pas d'attente) et les arcs de marche (pas de temps de déplacement, pas d'attente).

Par conséquent, le segment d'une ligne de transport en commun est un arc desservi par un véhicule à des intervalles donnés et l'usager du transport en commun attend que le lien soit desservi par un véhicule. Les arcs qui seront intégrés à une solution du modèle sont notés $\bar{A} \subseteq A$, où A représente l'ensemble des arcs et N représente l'ensemble des nœuds. Donc, la solution pour une destination s est le sous-graphe $b_s = (N, \bar{A})$. La demande de déplacements à partir des nœuds $i, i \in N$ jusqu'à la destination s est noté $\bar{g}_i$. Parmi les liens intégrés à la solution $\bar{A}$, à chaque nœud $i, i \in N$, un usager monte dans le premier véhicule qui dessert l'une ou l'autre des lignes dans $\bar{A}_i^+ (\bar{A} = \cup_i A_i^+)$. L'ensemble $\bar{A}_i^+$ correspond aux lignes qui seront choisies par l'usager pour former la route (ou les routes) entre i et s, dans une solution du modèle. À chaque arrêt i, il est opportun de faire référence à l'ensemble $\bar{A}_i^+$ comme étant un *ensemble de lignes attrayantes.*

Soit $W(\bar{A}_i^+)$, le temps d'attente prévu pour l'arrivée du premier véhicule desservant l'un ou l'autre des liens $a \in \bar{A}_i^+$, noté comme étant le temps d'attente combiné des liens

$a \in \bar{\mathcal{A}}_i^+$. Soit $P_a(\bar{\mathcal{A}}_i^+)$, la probabilité que le lien a sera la première ligne desservie parmi les liens $\bar{\mathcal{A}}_i^+$. Si une distribution exponentielle de l'intervalle entre les arrivées est admise, alors

$$W(\mathcal{A}_i^+) = \frac{1}{\sum\limits_{a \in \bar{\mathcal{A}}_i^+} f_a} \tag{67}$$

et

$$P(\mathcal{A}_i^+) = \frac{f_a}{\sum\limits_{a' \in \bar{\mathcal{A}}_i^+} f_a}, \; a \in \bar{A}_i^+, \tag{68}$$

où f_a est la fréquence du lien (ligne) a.

Puisque $\bar{\mathcal{A}}$ n'est pas connu à priori, le modèle de destination unique est formulé à l'aide de variables binaires $x_a, \; a \in \mathcal{A}$

$$x_a = \begin{cases} 0 & if \quad a \notin \bar{\mathcal{A}}, \\ 1 & if \quad a \in \bar{\mathcal{A}}. \end{cases}$$

Le modèle d'optimisation peut donc être formulé comme suit :

$$\text{Minimiser} \quad \sum_{a \in \mathcal{A}} s_a v_a + \sum_{i \in \mathcal{I}} \frac{V_i}{\sum\limits_{a \in \mathcal{A}_i^+} f_a x_a} \tag{69}$$

sous les contraintes

$$v_a = \frac{x_a f_a}{\sum\limits_{a' \in \mathcal{A}_i^+} f_a, x_a}, \quad a \in \mathcal{A}_i^+, \quad i \in \mathcal{N}, \tag{70}$$

$$V_i = \sum_{a \in \mathcal{A}_i^-} v_a + \bar{t}_i, \quad i \in \mathcal{N}, \tag{71}$$

$$V_i \geq 0, \quad i \in \mathcal{N}, \tag{72}$$

$$x_a \in \{0, 1\}, \quad a \in \mathcal{A}, \tag{73}$$

où s_a représente le coût des déplacements sur le lien a, et V_i représente le volume total au nœud i.

À première vue, le problème en 69–73 est un problème d'optimisation non-linéaire en nombres entiers. Heureusement, le problème peut être simplifié en le transformant en problème de programmation linéaire, grâce aux observations suivantes. Il est possible de remplacer 72 par les contraintes de non-négativité des volumes du lien $v_a \geq 0, a \in \mathcal{A}$, puisque $\sum\limits_{a \in \mathcal{A}_i^+} v_a = V_i, \; i \in \mathcal{N}$. En intégrant de nouvelles variables w_i, qui représentent le

temps d'attente total pour tous les déplacements au nœud i, $w_i = \frac{V_i}{\sum_{a \in \mathcal{A}_i^+} f_a x_a}$, $i \in \mathcal{N}$, nous obtenons le problème équivalent

$$\text{Minimiser} \quad \sum_{a \in \mathcal{A}} s_a v_a + \sum_{i \in \mathcal{N}} w_i \tag{74}$$

compte tenu de

$$v_a = x_a f_a w_i, \qquad a \in \mathcal{A}_i^+, i \in \mathcal{N}, \tag{75}$$

$$\sum_{a \in \mathcal{A}_i^+} v_a - \sum_{a \in \mathcal{A}_i^-} v_a = \bar{t}_i, \quad i \in \mathcal{N}, \tag{76}$$

$$v_a \geq 0, \quad a \in \mathcal{A}. \tag{77}$$

La fonction économique 74 est maintenant linéaire et les variables 0-1 sont seulement utilisées dans les contraintes 75, qui sont les seules contraintes non-linéaires. Il peut y avoir relaxation des contraintes en remplaçant 75 par

$$v_a f_a w_i, \quad a \in \mathcal{A}_i^+, \quad i \in \mathcal{N}, \tag{78}$$

ce qui produit le problème de programmation linéaire 74, 78, 76 et 77. Il est possible de démontrer, en utilisant les propriétés des points extrêmes que l'on trouve dans les solutions du modèle de programmation linéaire, que ce problème est équivalent à 74–77. Le problème dual du dernier programme linéaire est

$$\text{Maximiser} \quad \sum_{i \in \mathcal{N}} \bar{t}_i u_i \tag{79}$$

sujet à

$$u_j + s_a + \mu_a \geq u_i, \quad a \in \mathcal{A}, \tag{80}$$

$$\sum_{a \in \mathcal{A}_i^+} f_a \mu_a = 1, \quad i \in \mathcal{N}, \tag{81}$$

$$\mu_a \geq 0, \quad a \in \mathcal{A}, \tag{82}$$

où u_i, u_j sont les variables duales qui correspondent à 76 et μ_a sont les variables duales qui correspondent à 75.

Soit (v^*, w^*) et (u^*, μ^*), les solutions optimales des problèmes dual et primal. Les conditions faibles des écarts complémentaires sont :

$$(v_a^* - f_a w_i^*)\mu_a^* = 0, \quad a \in \mathcal{A}_i^+, \quad i \in \mathcal{N} \tag{83}$$

et

$$(u_i^* + s_a + \mu_a^* - u_i^*)v_a^* = 0, \quad a \in \mathcal{A}. \tag{84}$$

Dans les formulations primales et duales, le modèle de choix de route dans le transport en commun est semblable au modèle de choix du plus court chemin. Ce dernier correspond

au problème du plus court chemin lorsque l'attente est nulle sur tous les liens du réseau ; donc $f_a \to \infty$ et $w_i \to 0$. L'algorithme qui fournit la solution au modèle de choix de route pour le transport en commun est très semblable à l'algorithme de fixation d'étiquette utilisé pour calculer les plus courts chemins.

L'algorithme comprend deux parties. Dans la première passe, on calcule, en allant des nœuds de destination aux origines, les arcs sur lesquels il y a un flot, $\bar{A}^*$, et on calcule les temps de déplacement prévus u_i^* de chaque nœud $i, i \in \mathcal{N}$, aux nœuds de destination. Dans la deuxième passe, en allant des origines aux destinations, la demande est affectée aux arcs $a, a \in \bar{A}^*$. On écrit l'*algorithme de choix de route dans le transport en commun* comme suit (la convention $0 \infty = 1$ est utilisée), où les variables auxiliaires $f_i, i \in \mathcal{N}$ contiennent les fréquences combinées de tous les liens choisis au nœud i :

Première partie : Trouver $\bar{A}^*$

Étape 1.1. (Initialisation) $u_i \leftarrow \infty,\ i \in \mathcal{N} - \{s\}\ ;\ u_s \leftarrow 0$

$f_i \leftarrow 0,\ i \in \mathcal{N}$;

$\mathcal{S} \leftarrow A;\ \bar{A} \leftarrow \emptyset.$

Étape 1.2. (Choisir prochain lien) Si $\mathcal{S} = \emptyset$, STOP ; autrement trouver $a = (i, j) \in \mathcal{S}$ tel que

$$u_j + s_a \leq u_{j'} + s_{a'}; a' = (i', j') \in \mathcal{S}. \quad \leftarrow \mathcal{S} - \{a\}\,.$$

Étape 1.3. (Mise à jour des étiquettes des nœuds). Si $u_i \geq u_i + s_a$,

$$u_i \leftarrow \frac{f_i u_i + f_a(u_j + s_a)}{f_i + f_a},$$

$$f_i \leftarrow f_i + f_a;\ \ \bar{A} = \bar{A} + \{a\}\,,$$

Aller à l'étape 1.2.

Deuxième partie : Affecter la demande à $\bar{A}^*$

Étape 2.1. (Initialisation) $V_i \leftarrow \bar{t}_i,\ i \in \mathcal{N}.$

Étape 2.2. (Charger) Pour chaque lien $a \in \mathcal{A}$ en ordre décroissant de $(u_j + s_a)$, faire

Si $a \in \bar{A}$, alors $v_a \leftarrow \frac{f_a}{f_i} V_i$ et $V_j \leftarrow V_j + v_a$;

Sinon, $v_a \leftarrow 0$.

En utilisant les formulations primales et duales du modèle de choix de route dans le transport en commun, il est possible de prouver que l'algorithme trouve la solution de 74, 78, 76 et 77. L'algorithme est appliqué à toutes les destinations, tour à tour.

Il est possible d'appliquer l'algorithme à une version non-linéaire du problème, où les temps de déplacement sur les liens sont non pas constants, mais des fonctions continues $s_a(v_a), a \in \mathcal{A}$, des flots sur l'arc v_a. Le modèle qui en résulte peut être résolu par une adaptation de l'algorithme d'approximation linéaire. Pour en savoir davantage, veuillez consulter Spiess [38] et Spiess et Florian [39].

5 Planification stratégique de systèmes multimodaux

Les modèles et les méthodes présentés dans cette section ont une large portée : ils s'appliquent à des problèmes de planification stratégique d'envergure internationale, nationale et régionale, où les mouvements de plusieurs produits par les réseaux et services de plusieurs transporteurs sont considérés simultanément. Les principales questions portent sur l'évolution d'un système de transport donné et de sa réponse aux diverses transformations de son environnement : l'évolution de l'environnement socio-économique « local » ou international qui se traduit par des modifications des schémas et des volumes de production, de consommation et de commerce ; des changements aux politiques et aux lois existantes et l'instauration de nouvelles réglementations (par ex., les taxes liées à l'environnement dans certains pays européens) ; des modifications à l'infrastructure existante ; des fluctuations des prix de l'énergie ; des modifications aux conditions de travail ; des fusions de transporteurs ; l'avènement de nouvelles technologies, ainsi de suite. Ces enjeux font souvent partie intégrante des analyses coûts-bénéfices et des études comparatives sur des solutions de rechange en matière de politiques et d'investissement. Les organismes de planification et de réglementation aux divers paliers gouvernementaux s'intéressent tout particulièrement à ces questions, au même titre que les institutions financières internationales. Ces questions suscitent aussi un intérêt auprès des entreprises privées, notamment les sociétés engagées dans le financement des infrastructures de transport ou les compagnies qui planifient et exploitent la distribution des marchandises en ayant recours à plusieurs modes de transport.

Une méthodologie de planification stratégique « complète » représente les éléments fondamentaux d'un système de transport — demande, offre, mesures de la performance et critères de décision — et leurs interactions. Elle fournit les volumes de flot par produit et mode de transport, ainsi que les mesures de la performance associées, définis sur une représentation en réseau du système de transport. Elle vise à produire une simulation suffisamment précise du comportement global du système afin d'offrir une représentation juste de la situation actuelle et de servir d'outil d'analyse adéquat pour une large gamme de scénarios. Elle doit inclure des modèles qui peuvent être résolus et produire des résultats facilement accessibles. Cette démarche étant d'une portée très large, il est irréaliste de penser qu'une seule formulation ou procédure, mathématique ou autre, puisse englober tous les éléments pertinents, prendre en considération tous les enjeux importants et atteindre tous les objectifs.

Par conséquent, la méthodologie de planification se présente comme un ensemble de modèles et de procédures. Outre les outils de manipulation des données (par ex., collecte, fusion, mise à jour, validation, etc.) et les capacités d'analyse des résultats (par ex., coûts-bénéfices, incidences environnementales, politiques de consommation d'énergie, etc.), les principaux éléments de cette méthodologie sont :

1. la modélisation de l'offre, qui représente les modes de transport, l'infrastructure, les transporteurs, les services et les lignes ; les véhicules et les convois ; les terminaux et les installations intermodales ; les capacités et la congestion ; les mesures et les

critères en matière d'économie, de service et de performance ;

2. la modélisation de la demande, qui définit les produits ; identifie les producteurs, les expéditeurs et les intermédiaires ; représente la production et la consommation dans chaque zone, les volumes à transporter de zone à zone (région à région), ainsi que les choix de mode de transport ; ce volet traite aussi des relations de la demande et du choix de mode par rapport à la performance des politiques économiques et à la performance du système de transport ;
3. l'affectation des flots multiproduit (à partir du modèle de la demande) au réseau multimodal (la représentation de l'offre). Cette procédure simule le comportement global du système de transport et son résultat sert de base aux analyses stratégiques et aux activités de planification. Par conséquent, la méthodologie d'affectation doit être à la fois précise dans sa reproduction des situations actuelles et suffisamment générale pour produire des analyses solides fondées sur des données prévisionnelles.

La prédiction des flots de fret multiproduit sur un réseau multimodal est une composante importante de la science des transports et elle a soulevé beaucoup d'intérêt au cours des dernières années. Toutefois, peut-être en raison des difficultés et complexités inhérentes à de tels problèmes, on constate que l'étude des mouvements de fret à l'échelle nationale ou régionale n'a pas encore atteint sa pleine maturité, comparativement au transport des voyageurs pour lequel la prédiction des flots de voitures et du transport en commun a été largement étudiée et plusieurs résultats de recherche ont été transférés dans la pratique (Florian et Hearn [13] ; Cascetta [3]). Dans les pages qui suivent, nous faisons un rappel des méthodes les plus fréquemment employées pour la planification du transport de marchandises et passons brièvement en revue les références qui s'y rapportent.

5.1 Demande

La modélisation de la demande consiste à présenter une image des activités économiques d'un pays : production, consommation, importation et exportation de marchandises. Conçue à des fins de planification, elle a pour résultat une série de matrices de la demande spécifique d'un produit (ou d'un groupe de produits) qui indique les volumes à déplacer d'une région ou d'une zone à une autre. Ce processus est souvent complété par la modélisation du choix de mode, qui spécifie, pour chaque produit et combinaison origine-destination, sur quel ensemble d'infrastructure ou de services de transport la demande pourra être déplacée.

Un certain nombre de pays ont mis au point des modèles entrée/sortie de leur économie qui servent à déterminer la production de base et l'attraction des produits (Isard [23], Cascetta [3] et références incluses). Pour utiliser un modèle entrée/sortie, il est nécessaire de désagréger les données d'entrée et de sortie du modèle par région, puis de les désagréger encore en fonction de la subdivision zonale du modèle de planification national. Ce processus est complexe et il est habituellement réalisé dans un contexte d'analyse et de calcul

qui n'est pas nécessairement incorporé à celui employé pour la représentation de l'offre et le calcul des flots par produit. Lorsqu'un modèle entrée/sortie n'est pas disponible, on procède à la détermination initiale des matrices origine-destination au moyen des statistiques nationales sur la production, la consommation, les importations et les exportations que l'on conjugue à des études sectorielles destinées en compléter l'information manquante ou peu fiable. Ce processus peut se révéler fastidieux puisqu'il faut regrouper des données provenant de plusieurs sources qu'on recueille en utilisant différentes subdivisions géographiques ou définitions de produits non compatibles. Les résultats du modèle entrée/sortie désagrégé ou des procédures d'estimation ponctuelles servent au calcul initial des matrices origine-destination pour chaque produit, mais sans subdivision par mode.

Une seconde classe de modèles amplement étudiés pour la prédiction des flux interrégionaux de produits est formée par le modèle d'équilibre spatial et ses variantes (Friesz, Tobin et Harker [16], Harker et Friesz [20, 21] et Harker [18] ; voir aussi Florian et Hearn [13], Nagurney [32] et la section 2). Cette classe de modèles détermine simultanément les flux entre les régions productrices et consommatrices, ainsi que les prix de vente et d'achat qui satisfont les conditions de l'équilibre spatial. Autrement dit, un équilibre spatial est atteint à condition que, pour toutes les paires des régions d'offre et de demande présentant un flot de marchandises positif, le prix unitaire de l'offre additionné au coût unitaire du transport soit égal au prix unitaire de la demande ; la somme est plus grande que ce prix pour toutes les paires de régions sans échanges. On utilise généralement un réseau simple (graphe biparti) pour représenter le système de transport. Ces modèles s'appuient dans une large mesure sur les fonctions d'offre et de demande des producteurs et des consommateurs, respectivement, qui sont rarement disponibles et plutôt difficiles à calibrer. Il existe relativement peu d'applications de cette classe de modèles pour la détermination de la demande par produit. Les rares applications signalées dans la littérature se rapportent à des produits spécifiques qui ont une importance particulière tels que le pétrole brut, le charbon ou les produits laitiers.

5.2 Choix de mode

La définition du choix de mode peut être plutôt générale, par ex. correspondre aux mouvements de pétrole par navire et par pipeline ou, au contraire, être extrêmement spécifique et indiquer le chemin multimodal particulier pour un produit, un expéditeur ou une paire origine-destination donnée ou encore se situer quelque part entre les deux. Le niveau de détail de la spécification modale n'a pas besoin d'être la même pour tous les produits ou flots interzonaux. La spécification du choix de mode pour un produit donné peut être inférée à partir des données historiques et des études des expéditeurs ou elle peut résulter d'une description formelle et de travaux de modélisation (Winston [43]).

Les modèles d'utilité aléatoires, qui ont été développés et largement employés pour l'analyse et la planification des systèmes de transport de personnes, ont aussi été pro-

posés pour le transport de marchandises (Cascetta [3]), mais leur utilisation dans des applications réelles est rare. Ce phénomène est peut-être attribuable au grand nombre de chemins qu'il est nécessaire de générer et de stocker de façon explicite ainsi qu'à la difficulté d'exécuter cette tâche pour des données prévisionnelles. Aux niveaux agrégés, on a spécifié les choix de mode pour des flux particulièrement importants de produits en relevant de manière explicite les principales chaînes logistiques utilisées entre les paires des macrorégions.

5.3 Représentation de l'offre de transport et affectation de la demande

Une fois qu'on a créé des matrices modales origine-destination de quelque façon que ce soit, l'étape suivante consiste à les affecter au modèle de réseau (offre) en utilisant un mécanisme de choix d'itinéraire. Les résultats d'un tel modèle d'affectation – flots de produits et mesures de la performance – constituent une partie des données d'entrée pour la modélisation et l'analyse de la demande et des coûts-bénéfices. Le mécanisme d'affectation peut être fondé sur une application plus poussée des modèles d'utilité aléatoire au choix de chemins prédéfinis sur un réseau multimodal ou sur des modèles d'optimisation du réseau. Il est intéressant de noter que les attributs des chemins prédéfinis sont déterminés par l'état du réseau au moment de leur génération et ne sont pas influencés par les résultats de l'affectation. Ainsi, les conditions de congestion sont très difficiles à représenter. En outre, les modèles d'utilité et de choix doivent être calibrés et tous les chemins doivent être générés pour chaque scénario, ce qui est assez difficile à réaliser lorsqu'on utilise des données prévisionnelles.

Les modèles d'optimisation de réseau (voir la section 3) sont généralement considérés comme étant plus appropriés pour ce type de problèmes de planification. Ces formulations permettent la prédiction de flots multiproduit sur un réseau multimodal qui représente les installations de transport à un niveau de détail adéquat pour un pays ou une région, tout en étant relativement peu abstrait. La demande de services de transport est exogène et peut avoir pour origine un modèle entrée/sortie ou un modèle d'équilibre spatial, s'il y en a un de disponible, ou encore d'autres sources comme la demande observée ou l'ajustement de la demande observée antérieure. Le choix des sous-ensembles de modes permis pour chaque produit est exogènes et les acheminements intermodaux sont autorisés. À l'intérieur du choix de mode spécifié, le moteur d'optimisation (d'affectation) détermine les meilleurs chemins multimodaux pour chaque produit et paire origine-destination (en tenant compte des mesures de la performance spécifiées pour le réseau). Ces modèles peuvent aussi être intégrés dans les modèles économétriques de la demande. L'accent est mis sur une représentation adéquate du réseau et de ses divers modes de transport, les opérations de transfert intermodal correspondantes, les divers critères utilisés pour déterminer le mouvement des marchandises, les interactions et la concurrence relatives aux ressources limitées illustrées par la représentation des effets de la congestion et l'estimation de la distribution du trafic sur le système de transport associée que l'on utilisera

pour des études comparatives ou pour des analyses multipériode en temps discret.

Les études des années 1970 utilisaient des représentations plutôt simples des réseaux (par ex., Jones et Sharp [27] et Sharp [35]). Plusieurs études ont aussi essayé d'étendre les modèles d'équilibre spatial pour y inclure des représentations plus raffinées des réseaux et pour prendre en considération les effets de la congestion et les interactions entre les expéditeurs et les transporteurs. Friesz, Gottfried et Morlok [14] présentent un modèle séquentiel qui fait appel à deux représentations de réseaux : des réseaux distincts détaillés pour chaque transporteur et un réseau global considéré du point de vue de l'expéditeur. Sur chaque réseau de transporteur, les marchandises sont transportées au moindre coût total. Sur le réseau considéré du point de vue de l'expéditeur, on utilise des principes d'équilibre usager du trafic pour déterminer les transporteurs que les expéditeurs choisissent pour déplacer leur trafic. Cette approche a permis d'obtenir d'assez bons résultats dans l'étude de la logistique des produits lorsqu'un nombre très limité d'expéditeurs et de transporteurs interagissent et déterminent fortement le comportement du système. Le marché du charbon entre les services publics d'électricité aux États-Unis et leurs fournisseurs dans les pays exportateurs en est un exemple typique. Friesz et Harker [15], Harker et Friesz [20, 21], Harker [18, 19], Hurley et Petersen [22] et Fernández, de Cea et Soto [12] ont présenté des formulations plus élaborées. Cette piste de recherche n'a toutefois pas encore produit de modèles et outils pratiques de planification, principalement parce que les formulations deviennent trop vastes et complexes lorsqu'on les applique à des situations réalistes.

Les méthodologies qui ont produit des outils de planification transférés avec succès dans la pratique utilisent de représentations sophistiquées des réseaux et services de transport et simulent le comportement global du système à l'aide de modèles d'optimisation de réseau (Guélat, Florian et Crainic [17], Crainic, Florian et Léal [8], Crainic, Florian et Larin [7], Crainic *et al.* [6], [4], [5], Jourquin et Beuthe [28]). Le cadre de modélisation que nous présentons est fondé sur les travaux de Guélat, Florian et Crainic [17].

La formulation ne considère pas les expéditeurs et les transporteurs comme des acteurs distincts dans les décisions prises relativement à l'envoi des marchandises. Le niveau d'agrégation, approprié pour la planification stratégique des flots de fret, donne lieu à des origines et des destinations qui correspondent à des zones géographiques relativement vastes et à la spécification de l'offre et de la demande représentant, pour chacun des produits considérés, les volumes totaux générés par tous les expéditeurs individuels. De plus, la demande dans les études stratégiques sur le transport des marchandises est souvent déterminée à partir de sources de données (statistiques nationales de flux de fret, modèles économiques entrée/sortie) qui permettent d'établir le mode utilisé, mais qui ne contiennent pas d'information sur les expéditeurs individuels. On présume donc que le comportement de l'expéditeur se reflète dans les matrices des produits origine-destination et la spécification du choix de mode correspondant.

Le cadre de modélisation est celui d'un réseau multimodal, constitué de modes, de nœuds, de liens et de transferts intermodaux, sur lequel de multiples produits doivent

être déplacés par des véhicules et des convois spécifiques entre des points d'origine et de destination donnés. Ici, un mode est un moyen de transport qui possède ses propres caractéristiques, telles que le type de véhicule et la capacité, ainsi que ses mesures de coût spécifiques. Selon la portée et le niveau de détail de l'étude stratégique, un mode peut représenter un transporteur ou une portion de son réseau correspondant à un service de transport particulier, à une agrégation de plusieurs réseaux de transporteurs ou à des infrastructures de transport spécifiques comme des réseaux d'autoroutes ou des ports.

Le réseau est constitué de nœuds $\mathcal{N}$, de liens $\mathcal{A}$, de modes $\mathcal{M}$ et de transferts $\mathcal{T}$ qui représentent tous les mouvements physiques possibles sur l'infrastructure disponible. Pour reproduire les caractéristiques modales du transport, un lien $a \mathcal{A}$ est défini sous la forme d'un triplet (i, m, j), où $i \in \mathcal{N}$ est le nœud d'origine, $j \in \mathcal{N}$ est le nœud de destination et $m \in \mathcal{M}$ est le mode alloué sur l'arc. Des liens parallèles sont utilisés pour représenter des situations où il y a plus d'un mode disponible pour transporter les marchandises entre deux nœuds adjacents. Cette représentation en réseau permet une détermination facile du flots de marchandises par mode, ainsi que les diverses fonctions de coût (par ex., coûts d'exploitation, retards, consommation d'énergie, émissions, bruits, risques, etc.) par produit et par mode. Pour modéliser les acheminements intermodaux, il faut tenir compte des transferts de mode à certains nœuds du réseau et calculer les coûts et les retards qui y sont associés. Les transferts intermodaux $t \in \mathcal{T}$ à un nœud du réseau sont modélisés sous la forme de mouvements alloués de lien à lien et, par conséquent, de mode à mode. Un chemin dans ce réseau consiste alors en une séquence de liens orientés d'un mode, en un transfert possible à un autre mode, en une séquence de liens orientés d'un second mode et ainsi de suite. Par conséquent, un transfert appartient à un chemin si les deux arcs qui le définissent appartiennent au chemin. Cette représentation en réseau tient compte de la restriction des flots de certains produits à des sous-ensembles de modes (par ex. le minerai de fer ne peut être expédié que par rail ou par navire) afin de reproduire les décisions particulières en matière de choix de mode ou les restrictions qui surviennent dans l'exploitation de réseaux de transport de marchandises et d'installations de transbordement.

On entend par produit toute marchandise ou tout passager (ou collection de produits similaires) qui génère un flot de liens. Chaque produit $p \in \mathcal{P}$ transporté sur le réseau multimodal est expédié depuis certaines origines $o \in \mathcal{N}$ jusqu'à certaines destinations $d \in \mathcal{N}$ du réseau. La demande pour chaque produit pour toutes les paires origine-destination est exogène et est spécifiée par un ensemble de matrices O-D. Le choix de mode pour chaque produit est lui aussi exogène et on l'indique en définissant pour chaque matrice O-D un sous-ensemble de modes alloués pour le transport de la demande correspondante. Soit $g^{m(p)}$ une matrice de demande associée au produit $p \in \mathcal{P}$ où $m(p) \subseteq \mathcal{M}$ est le sous-ensemble des modes qui peuvent être utilisés pour déplacer cette portion particulière du produit p.

Les flots du produit $p \in \mathcal{P}$ sur le réseau multimodal sont les variables de décision du modèle. Les flots sur les liens $a \in \mathcal{A}$ sont représentés par v_a^p et les flots sur les transferts $t \in \mathcal{T}$ sont désignés par v_t^p ; v représente le vecteur de tous les flots de produits. Les

fonctions de coût sont associées aux liens et aux transferts du réseau. Pour le produit p, les fonctions de coût moyen respectives $s_a^p(v)$ et $s_t^p(v)$ dépendent du volume de marchandises transportées. Donc, le coût total du produit p sur l'arc a est $s_a^p(v)v_a^p$ et il correspond à $s_t^p(v)v_t^p$ sur le transfert t. Le coût total sur le réseau multimodal est la fonction F, qui sera minimisée sur l'ensemble de volumes des flots sur les liens et les transferts qui satisfont les contraintes de conservation et de non-négativité des flots :

$$F = \sum_{p \in P} \left(\sum_{a \in \mathcal{A}} s_a^p(v) v_a^p + \sum_{t \in T} s_t^p(v) v_t^p \right). \tag{85}$$

Soit $\mathcal{L}_{od}^{m(p)}$ qui représente l'ensemble des chemins qui, pour le produit p, vont de l'origine o à la destination d en utilisant seulement les modes dans $m(p)$. La formulation « chemin » des équations de conservation des flots est donc :

$$\sum_{l \in \mathcal{L}_{od}^{m(p)}} h_l = g_{od}^{m(p)} \qquad o,\ d \in \mathcal{N},\ p \in \mathcal{P},\ m(p) \subseteq \mathcal{M}, \tag{86}$$

où h_l est le flot sur le chemin $l \in \mathcal{L}_{od}^{m(p)}$. Ces contraintes précisent que le flot total déplacé sur tous les chemins pouvant être utilisés pour transporter le produit p doit être égal à la demande pour ce produit. Les contraintes de non-négativité sont :

$$h_l \geq 0, \qquad l \in \mathcal{L}_{od}^{m(p)},\ o,\ d \in \mathcal{N},\ p \in \mathcal{P},\ m(p) \subseteq \mathcal{M}. \tag{87}$$

La relation entre les flots sur les arcs et les flots sur les chemins est $v_a^p = \sum_{l \in \mathcal{L}^p} \delta_{al} h_l,\ a \in \mathcal{A},\ p \in \mathcal{P}$, où $\mathcal{L}^p$ est l'ensemble de tous les chemins qui peuvent être utilisés par le produit p, et $\delta_{al} = 1$ si $a \in l$ (et 0, autrement) est la fonction indicatrice qui identifie les arcs d'un chemin particulier. De même, les flots sur les transferts sont $v_t^p = \sum_{l \in \mathcal{L}^p} \delta_{tl} h_l,\ t \in \mathcal{T},\ p \in \mathcal{P}$, où $\delta_{al} = 1$ si $t \in l$ (et 0, autrement). Le modèle d'affectation multimodal multiproduit consiste donc à minimiser 85, sous réserve des contraintes 86 et 87. C'est un modèle d'optimisation système et le principe d'optimalité garantit que dans la distribution finale des flot, pour chaque produit, matrice de demande et paire origine-destination, tous les chemins présentant des flot positifs auront le même coût marginal (moindre que sur les autres chemins). L'algorithme mis au point pour ce problème exploite la décomposition naturelle par produit et résulte en une procédure de type Gauss-Seidel qui permet de résoudre des problèmes à grande échelle dans des temps de calcul raisonnables (Guélat, Florian et Crainic [17] ; Section 3.2.2).

Ce modèle de réseau permet une représentation détaillée de l'infrastructure, des installations et des services ainsi que l'affectation simultanée de produits multiples sur des modes multiples. Le trafic des véhicules et des convois sur les liens (et les transferts) du réseau est déterminé à partir des flots de produits affectés et il est utilisé pour évaluer les conditions de congestion et calculer les coûts. Les capacités sont considérées au moyen des fonctions de congestion ou de pénalité. Ainsi, le modèle indique la concurrence entre

les produits par rapport à la capacité de services disponibles, une caractéristique d'une grande pertinence lorsqu'on considère des scénarios de rechange pour l'accroissement de la capacité du réseau. Il rend compte de la spécification et de la combinaison d'une grande diversité de mesures de la performance et de critères d'affectation, y compris des fonctions de type optimisation usager quand la nature d'un produit particulier l'exige. En outre, le modèle est suffisamment flexible pour représenter l'infrastructure d'un seul transporteur.

6 Conclusions

Nous avons présenté plusieurs modèles d'optimisation qui sont utilisés fréquemment dans l'exercice de la planification du transport, tant de passagers que de marchandises. Nous avons vu, en particulier, le modèle d'interaction spatiale, des variantes du modèle d'équilibre de réseau et un modèle de choix de route dans le transport en commun. Les algorithmes de solution associés à chacune des catégories de modèles ont également été évoqués. Cette présentation illustre la richesse de l'application de la recherche opérationnelle au domaine de la planification du transport.

7 Remerciements

Ce projet a été financé en partie par le Conseil de recherches en sciences naturelles et en génie du Canada, par le biais du programme de subventions de recherche *Subventions à la découverte*.

Références

[1] Andersson, P.A. On the Convergence of Iterative Methods for the Distribution Balancing Problem. *Transportation Research B : Methodological*, 15 :173–217, 1981.

[2] Bregman, L.M. The Relaxation Method for Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming. *URSS Computational Mathematics and Mathematical Physics*, 7 :200–217, 1967.

[3] Cascetta, E. *Transportation Systems Engineering : Theory and Methods.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

[4] Crainic, T.G., Dufour, G., Florian, M., and Larin, D. Path Analysis in STAN. In C. Zopounidis and D. Despotis, editors, *Proceedings 5th International Conference of the Decision Sciences Institute*, pages 2060–2064. New Technologies Publications, Athens, Greece, 1999.

[5] Crainic, T.G., Dufour, G., Florian, M., and Larin, D. Path Recovery/Reconstruction and Applications in Nonlinear Multimodal Multicommodity Networks. In Gendreau, M. and Marcotte, P., editors, *Transportation and Network Analysis : Current Trends Miscellanea in Honor of Michael Florian*, pages 2060–2064. Kluwer Academic Publishers, Norwell, MA, 2002.

[6] Crainic, T.G., Florian, M., Guélat, J., and Spiess, H. Strategic Planning of Freight Transportation : STAN, An Interactive-Graphic System. *Transportation Research Record*, 1283 :97–124, 1990.

[7] Crainic, T.G., Florian, M., and Larin, D. STAN : New Developments. In Al S. Khade and R. Brown, editors, *Proceedings of the 23rd Annual Meeting of the Western Decision Sciences Institute*, pages 493–498. School of Business Administration, California State University, Stanislaus, CA, 1994.

[8] Crainic, T.G., Florian, M., and Léal, J.-E. A Model for the Strategic Planning of National Freight Transportation by Rail. *Transportation Science*, 24(1) :1–24, 1990.

[9] Demins, W.E. and Stephan, F.F. On a Least Squares Adjustment of a Sampled Frequency Yable when the Expected Marginal Totals are Known. *Annals of Mathematical Statistics*, 11 :427–444, 1940.

[10] S. Erlander, S. Nguyen, and N. Stewart. On the Calibration of the Combined Distribution-Assignment Model. *Transportation Research B : Methodological*, 13 :259–267, 1979.

[11] Evans, S.P. and Kirby, H.R. A Three-dimensional Furness Procedure for Calibrating Gravity Models. *Transportation Research*, 8 :105–122, 1974.

[12] Fernández, J.E., de Cea Ch., J., and Soto O., A. A Multi-modal Supply-Demand Equilibrium Model for Predicting Intercity Freight Flows. *Transportation Research B : Methodological*, 37 :615–640, 2003.

[13] Florian, M. and Hearn, D. Network Equilibrium Models and Algorithms. In M. Ball, Magnanti, T.L., Monma, C.L., and Nemhauser, G.L., editors, *Network Routing*, volume 8 of *Handbooks in Operations Research and Management Science*, pages 485–550. North-Holland, Amsterdam, 1995.

[14] Friesz, T.L., Gottfried, J.A., and Morlok, E.K. A Sequential Shipper-Carrier Network Model for Predicting Freight Flows. *Transportation Science*, 20 :80–91, 1986.

[15] Friesz, T.L. and Harker, P.T. Freight Network Equilibrium : A Review of the State of the Art. In A.F. Daughety, editor, *Analytical Studies in Transport Economics*, chapter 7. Cambridge University Press, Cambridge, 1985.

[16] Friesz, T.L., Tobin, R.L., and Harker, P.T. Predictive Intercity Freight Network Models. *Transportation Research A : Policy and Practice*, 17A :409–417, 1983.

[17] Guélat, J., Florian, M., and Crainic, T.G. A Multimode Multiproduct Network Assignment Model for Strategic Planning of Freight Flows. *Transportation Science*, 24(1) :25–39, 1990.

[18] Harker, P.T. *Predicting Intercity Freight Flows*. VNU Science Press, Utrech, The Netherlands, 1987.

[19] Harker, P.T. Issues and Models for Planning and Regulating Freight Transportation Systems. In L. Bianco and A. La Bella, editors, *Freight Transport Planning and Logistics*, pages 374–408. Springer-Verlag, Berlin, 1988.

[20] Harker, P.T. and Friesz, T.L. Prediction of Intercity Freight Flows I : Theory. *Transportation Research B : Methodological*, 20B(2) :139–153, 1986.

[21] Harker, P.T. and Friesz, T.L. Prediction of Intercity Freight Flows II : Mathematical Formulations. *Transportation Research B : Methodological*, 20B(2) :155–174, 1986.

[22] Hurley, W.J. and Petersen, E.R. Nonlinear Tariffs and Freight Network Equilibrium. *Transportation Science*, 28(3) :236–245, 1994.

[23] Isard, W. Interregional and Regional Input-Output Analysis : A Model of a Space-Economy. *The Review of Economics and Statistics*, 33 :318–328, 1951.

[24] Jaynes, E. Information Theory and Statistical Mechanics. *Physical Review*, 106 :171–190, 1957.

[25] Jaynes, E. Information Theory and Statistical Mechanics. *Physical Review*, 106 :620–630, 1957.

[26] Jefferson, T.R. and Scott, C.H. The Analysis of Entropy Models with Equality and Inequality Constraints. *Transportation Research B : Methodological*, 13 :123–132, 1979.

[27] Jones, P.S. and Sharp, G.P. Multi-Mode Intercity Freight Transportation Planning for Underdeveloped Regions. In *Proceedings of the 18th Meeting of the Transportation Research Forum*, pages 523–531, 1977.

[28] Jourquin, B. and Beuthe, M. Transportation Policy Analysis with a Geographic Information System : The Virtual Network of Freight Transportation in Europe. *Transportation Research C : Emerging Technologies*, 4(6) :359–371, 1996.

[29] J. Kruithof. Calculation of Telephone Trafic. *De Ingenier*, 52 :E15–E25, 1937. en flamand.

[30] Kullback, S. *Information Theory and Statistics*. Wiley, New York, 1959.

[31] Lamond, B. and Stewart, N.F. Bregman's Balancing Method. *Transportation Research B : Methodological*, 15 :239-248, 1981.

[32] Nagurney, A. *Network Economics : A Variational Inequality Approach*. Kluwer Academic Publishers, Boston, second edition, 1999.

[33] Patricksson, P. *The Traffic Assignment Problem : Models and Methods*. VNU Science Press, Utrecht, 1993.

[34] Robillard, P. and Stewart, N.F. Iterative Numerical Methods for Trip Distribution Problems. *Transportation Research*, 8 :575–582, 1974.

[35] Sharp, G.P. A Multi-Commodity, Intermodal Transportation Model. In *Proceedings of the 20th Meeting of the Transportation Research Forum*, pages 399–407, 1979.

[36] Sheffi, Y. *Urban Transportatin Networks. Equilibrium Analysis with Mathematical Programming*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[37] Snickars, F. and Weibull, J.W. A Minimum Information Principle, Theory and Practice. *Regional Science and Urban Economics*, 7 :137–168, 1990.

[38] H. Spiess. *Contributions à la théorie et aux outils de planification des réseaux de transport urbains.* PhD thesis, Université de Montréal, Montréal, QC, Canada, 1984.

[39] H. Spiess and M. Florian. Optimal Strategies : A New Assignment Model for Transit Networks. *Transportation Research B : Methodological*, 23 :83–102, 1981.

[40] Wardrop, J.G. Some Theoretical Aspects of Road Traffic Research. In *Proceedings Institution of Civil Engineers, Part II*, pages 325–378, 1952.

[41] Wilson, A.G. A Statistical Theory of Spatial Distribution Models. *Transportation Research*, 1 :253–269, 1967.

[42] Wilson, A.G. *Entropy in Urban and Regional Modelling.* Pion, Londres, 1970.

[43] Winston, C. The Demand for Freight Transportation : Models and Applications. *Transportation Research A : Policy and Practice*, 17 :419–427, 1983.